DOCUMENT RESUME

ED 183 520 SP 015 615

AUTHOR Johnson, Charles E.: And Others
TITLE An Investigation of the Legal Validity of the Teacher Performance Assessment Instruments: A Summary.
INSTITUTION Georgia Univ., Athens. Coll. of Education.
SPONS AGENCY Georgia State Dept. of Education, Atlanta. Div. of Program and Staff Development.
PUB DATE 78
NOTE 56p.; For related documents, see SP 015 610, SP 015 612-613, SP 015 617.

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Beginning Teachers: *Court Litigation: *Due Process: *Equal Protection: *Evaluation Methods: *Occupational Tests: *Teacher Certification: Teacher Employment

ABSTRACT

Investigations into the legal validity of the content and procedures proposed in the Teacher Performance Assessment Instruments (TPAI) and allied materials for use in certifying Georgia's beginning teachers are reported. TPAI was designed to determine how well teachers can demonstrate certain minimum teaching skills considered essential to effective teaching. The report focuses upon a survey of reference materials requisite to a preliminary understanding of potential legal problems arising when the proposed assessment system and materials are eventually implemented officially by the State Department of Education. Thirty-two court cases dealing with due process and equal protection in reference to teacher assessment are cited and discussed in tabular format. A summary of findings from this survey of litigation is presented. (JD)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

ED183520

nov
'78
41p

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
A. Moughon

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

SP015615

TEACHER ASSESSMENT PROJECT
COLLEGE OF EDUCATION
UNIVERSITY OF GEORGIA
ATHENS 30602

# An Investigation of the Legal Validity of the *Teacher Performance Assessment Instruments:* A Summary

Charles E. Johnson

Pearce T. Adams

Kathleen A. Miller

Chad D. Ellett

William Capie

1978

College of Education
University of Georgia
Athens, Georgia 30602

The materials contained herein were prepared under a contract with the Georgia Department of Education, Division of Program and Staff Development, Office of Instructional Services, Atlanta, Georgia. Funds to conduct the research and development activities described in the contract were provided by the Board of Education of the State of Georgia and supplemented with contributions, facilities, and staff services by the University of Georgia.

University of Georgia
Teacher Assessment Project Staff

Director

Charles E. Johnson, Ed.D.

Associate Directors

William Capie, Ed.D.

Chad D. Ellett, Ph.D.

Coordinators

Pearce T. Adams, M.Ed.

S. Joy Anderson, M.Ed.

Thomas J. Arnold, M.Ed.

J. Anthony Masters, M.Ed.

Catherine E. Garrison

State Department of Education

Dr. Charles McDaniel
Superintendent of Schools

Dr. H. Titus Singletary
Associate Superintendent of Schools

Dr. J. William Leach
Director of Division of Staff Development

Dr. Mary L. Mitchell
Assistant Director of Staff Development

Lester M. Solomon
Coordinator for Teacher Education and Staff Development

John L. Hooper
Consultant for Performance Based Certification

The Georgia Teacher Education Council
Student Teacher--Beginning Teacher Ad Hoc Committee

Dr. James O. Miller, Chairperson
Dean of Education
Emory University
Atlanta

Dr. Bobby Jones
Coordinator of Student Teaching
Mercer University
Macon

Dr. Fred Broder, Representative
Georgia Association of Educators
Decatur

Dr. Joseph A. Murphy, Director
Performance Based Certification with Supportive Supervision Project
Clarkston

Dr. Therry Deal
Professor of Home Economics
Georgia College
Milledgeville

Dr. George L. O'Kelly, Jr.
Professor of Education
University of Georgia
Athens

Dr. Robert Driscoll, Dean
School of Education
Kennesaw College
Kennesaw

Miss Jacquelyn D. Ray, Teacher
Clarkston High School
Clarkston

Dr. Susan Fuller, Director
Griffin CESA Project
Griffin

Mr. Null A. Tucker
PBC/SS Project
Clarkston

Dr. Jacob T. Hunt
Professor of Education
University of Georgia
Athens

Mrs. Betty W. Whisnant
Harris County Board of Education
Hamilton

## Purpose

This is a report of a study designed to investigate the legal validity of the content and procedures proposed in the *Teacher Performance Assessment Instruments* (TPAI) and allied materials for use in certifying Georgia's beginning teachers. It focuses upon a survey of the reference materials requisite to a preliminary understanding of potential legal problems arising when the proposed assessment system and materials are eventually implemented officially by the State Department of Education (SDE).

Information generated by this study will be used by the Teacher Assessment Project (TAP) to refine the instruments and recommended procedures in such a manner that they will reflect legally acceptable approaches to teacher assessment for certification purposes.

## Background

TAP is a research organization housed at the University of Georgia and supported primarily by the State Board of Education through the Georgia State Department of Education. The sole function of TAP is to develop teacher performance instruments and procedures which can be implemented by the SDE in its newly mandated beginning teacher certification system.

To understand the role of performance assessment in Georgia's teacher certification program, one must be cognizant of the main differences between the previous certification program and the one which is currently becoming effective. There are a few clearly identifiable

differences. One is that in the future the initial certificate which is normally provided upon graduation from college will be valid for only three years, and it will be nonrenewable. Another is that, in the future, in order to obtain a valid *renewable* certificate, the beginning teacher must, within a three-year period or less after initial certification, have demonstrated competent teaching performance. The third difference is that there are two sets of instruments which were not used previously --the TPAI and the *Teacher Certification Tests* (TCT). The latter are sets of criterion-referenced examinations relating to teaching specialties.

In most cases, the TCT are administered to the prospective candidate during the candidate's senior year in college. They are used to determine whether the student has met certain minimum standards of knowledge and understanding in that teacher's teaching field.

To avoid erroneous assumptions, it should be pointed out that none of the previous certification criteria have been eliminated from the newly proposed requirements. Students must still obtain a college degree and take courses required by state-approved programs for teachers. There has been no attempt to reduce the number of credit hours required, the nature of the requirements for a major or minor, or the extent to which students study the theory and practice of education. This program simply adds teacher performance assessment standards that were not previously a part of the teacher certification requirements.

The first drafts of the TPAI were designed, developed, and piloted by TAP during the period beginning April 1976 and ending June 1978. These instruments assess a teacher's performance in four major areas:

*Teaching Plans and Materials, Classroom Procedures, Interpersonal Skills and Professional Standards*

The TPAI consist of a set of 69 indicators, each composed of five behavior descriptors, representing various teaching competencies. For beginning teachers, it has been proposed that each beginning teacher undergo two successful assessments within a three-year period. For each assessment, data will be provided by a trained data collector (a data source person external to the teacher's school), an administrator, and a master teacher from within the teacher's subject area. For student teachers the proposed procedures are similar, but different because their responsibilities and time commitment are limited.

At present only data collectors are being given the comprehensive training program. Ultimately it may be that all assessors will be equally well trained. Should the beginning teacher being assessed fail to meet the minimum standards of the assessment during the first year, it is proposed that this person would have two additional years in which to demonstrate the minimum performance standards.

Content of the TPAI reflects the findings of an initial study* soliciting the opinion of 4,500 Georgia educators as to what constitutes effective teacher performance. Of a total of 52 competencies presented to these educators, 20 were considered essential, regardless of the subject or grade level taught. These 20 generic competencies formed the foundation for development of the TPAI.

---

*Charles E. Johnson, et al., *Identifying and Verifying Generic Teacher Competencies* (Athens, Georgia: Teacher Assessment Project, 1978).

Other studies were conducted during the initial piloting period (Fall 1977 through Spring 1978) to provide input for revision of the TPAI for clarity, validity, and reliability prior to publication of a revised edition in the Spring of 1978. The revised TPAI will be used in a second piloting which began in the Summer of 1978 and extends into 1979.

Validity of the TPAI has been preliminarily established in a variety of ways. Content validity characteristics have been determined through statewide verification of competencies, indicators, and studies utilizing experts' opinions. Concurrent or "discriminant" validity has been examined by comparing teachers with various amounts of teaching experience, with the finding that more experienced teachers tend to score higher on the instruments. Validity has been examined through factor analysis of the competencies and indicators and by correlating TPAI assessments with the *Purdue Observation Rating Scale*. Criterion-related validity has been researched by comparing elementary and secondary teachers' TPAI scores with a variety of criterion variables such as: student achievement on standardized tests, student learning on teacher-made tests, observations of student "on-task" behavior, and measures of students' perceptions of characteristics of the school climate/learning environment.

Preliminary investigations have provided much information about the reliability of ratings. Investigations conducted via videotape and other standardized means have shown reliability estimates in the range of .70 to .90 in most instances. Observers in the classroom agree on the same scale point about 70% of the time. Adjacent scale points are chosen in

more than 90% of these instances. Currently, field-testing should reflect improvements made in the instruments and their use.

Evidence derived from these preliminary investigations is encouraging in the sense that the TPAI have demonstrated relationships with the various criterion measures used. Additional investigations with improved research designs and larger samples will be conducted to verify the results of the preliminary research activities.

## Procedure and Rationale

A variety of different resources were consulted in conducting the legal feasibility study. An initial computer scan of the Educational Resources Information Center's (ERIC) system provided minimally helpful information. The focus of the study then shifted to the resources of the University of Georgia law library. Various digests and encyclopedias were consulted to develop a list of court cases addressing teacher certification specifically and employment testing in general. Investigation of these cases produced not only valuable information in their own right, but often referenced additional pertinent material such as law reviews, books and articles.

Examination of court cases also led to an investigation of the relevant statutes upon which litigation is based, as well as to applicable guidelines and standards for employment testing. These primary materials included relevant sections of the Georgia Code; the Thirteenth and Fourteenth Amendments to the U.S. Constitution; Title VII of the 1964 Civil Rights Act and its amendment, the Equal Employment Opportunity Act of 1972; the Equal Employment Opportunity Commission Guidelines;

The Federal Executive Agency Guidelines; and the American Psychological Association's *Standards for Educational and Psychological Tests.*

## Findings

The findings of this study are too voluminous to be included here, but can be found in their entirety in a supplemental volume: Johnson, Charles E., et al. *Summary of Court Cases Pertinent to Performance-Based Certification in Georgia* (In-House Draft of a Working Paper). Athens, Georgia: 1978.

For the purposes of the present paper, findings have been encapsulated and presented in tabular fashion on the following pages. The court cases cited have been classified under two broad headings--those dealing with due process and those dealing with equal protection. These two areas have been further subdivided according to whether the complaint made was upheld or dismissed. Information on the date of the litigation, an abbreviated case name, a description of the issues involved, and the court's decision is presented in each instance.

The Selected Bibliography merely highlights the most important aspects of this research. A more extensive bibliography can be found in *Summary of Court Cases, supra.*

# TABLES

Table 1: Complaints Dealing with Due Process that Were Dismissed by the Courts

| Date | Case | Description | Court Decision |
|---|---|---|---|
| 1903 | Meffert | Medical license was revoked on the grounds of gross immorality. Plaintiff claimed that the license could not be revoked unless charge related to the practice of medicine. | Kansas Supreme Court held that the Board was within its discretion, and that unless abuse of discretion could be proven, the decision was conclusive. |
| 1907 | Stone | Teacher's license was revoked on the grounds of incompetence, immorality, cruelty, and general neglect. Teacher charged that the Superintendent exceeded his authority and violated his due process. | If a Superintendent acted within statutory provisions, he was within his jurisdiction and courts would not intervene unless discretion was abused. |
| NOTE: | In *Stone* the court said that in accepting a license, the licensee agrees to abide by the conditions set forth by the licensing statute. | | |
| 1925 | Marrs | Superintendent revoked a certificate of a teacher who participated in a scheme to issue fraudulent certificates. | Court upheld the dismissal and said that due process was not denied. |
| NOTE: | In *Marrs* the court said that a teaching certificate is not a contract protected by due process clauses of the state or federal constitutions. | | |
| 1952 | Hodge | Teacher requested that his certificate, revoked for 12 violations of State Law, be reissued on grounds statutes did not specify conduct of hearing, charges were unproven, and Board exceeded its authority. | Complaint was upheld by District Court, but overturned by Oklahoma Supreme Court. Board had authority to revoke in face of violation of state law even though procedural rules were unspecified. |

Table 1: Continued

| Date | Case | Description | Court Decision |
|---|---|---|---|
| 1971 | Guthrie | Teacher filed class action suit challenging a state policy requiring additional teaching credit for renewal of certificate. | North Carolina Supreme Court found rule was neither unreasonable nor arbitrary in attempt to improve instruction. No violation of due process. |
| 1976 | Washington | Class action claiming violation of due process and equal protection by discriminatory impact of a written promotion test on blacks. | Supreme Court held that the test did not deny due process or equal protection and that to prove the test unconstitutional, discriminatory intent as well as adverse impact must be shown. The positive relationship of the test to training school performance was sufficient to validate. |
| 1977 | Beachamp | Teacher challenged nonrenewal of contract due to principal's rating of incompetence on unvalidated evaluation as denying due process and equal protention. | Court found nonrenewal justified on other grounds and rejected due process claim. |

Table 2: Complaints Dealing with Due Process that Were Upheld by the Courts

| Date | Case | Description | Court Decision |
|---|---|---|---|
| 1943 | Matteson | Mandamus proceeding for renewal of a tenured teacher's certificate and restoration of job. | Nonrenewal of certificate and subsequent dismissal from job without bringing charges or providing hearing contravened established procedures. Mandamus granted. |
| 1971 | Baker | Discrimination action against use of an unvalidated NTE cutoff score for teacher employment, coinciding with court-ordered integration, denying due process and equal protection. | Use of NTE score not shown job-related found to have discriminatory intent and effect. |
| 1973 | Brito | Civil rights action charging employment evaluation test of on-the-job performance discriminated against Spanish workers. | The unvalidated evaluation disproportionately affected Spanish workers and was defective by EEOC Guidelines: it was subjective, not conducted under controlled and standardized conditions, and was not shown to be job-related. Enjoined. |

Table 2: Continued

| Date | Case | Description | Court Decision |
|---|---|---|---|
| 1974 | Walston | Challenge by black applicants to use of NTE for teacher selection in a recently desegregated school system as violating due process and equal protection. | NTE inequitably administered and used in discriminatory manner not conforming to designer's intent without proper job analyses and validation studies. NTE enjoined, teachers dismissed because of NTE ordered reinstated with backpay, and those dismissed for cause re-evaluated. |
| 1974 | Longarzo | Proceeding to vacate unsatisfactory teacher's rating due to denial of due process. | Failure of New York City Assistant Superintendent to review plaintiff's work before approving principal's unsatisfactory rating contravened policy and denied him a fair hearing without due process. Rating annulled. |
| 1976 | G.A.E. | Civil rights action to enjoin arbitrary use of NTE for certification due to discriminatory impact. | No rational relationship shown between NTE and Georgia's stated purpose for using it for certification. NTE used in a manner contrary to designer's intent without proper validation. Cutoff score arbitrary and violated equal protection. Enjoined. |

Table 3: Complaints Dealing with Equal Protection that Were Dismissed by the Courts

| Date | Case | Description | Court Decision |
|---|---|---|---|
| 1968 | U.S. v. Porter | Action by U.S. against discriminatory employment practices and use of general aptitude tests for transfer, claiming lack of professional validation constituted discrimination. | Court rejected government claim, holding no evidence of discriminatory effect shown. Although tests were not professionally validated, essential steps were taken. Use of tests upheld. |
| 1971 | Allen | Class action by black policemen against discriminatory employment practices and unvalidated sergeant's promotion exam. | Exam was ruled job-related and court refused to require validation. Quota promotions for blacks denied. |
| 1972 | Lee | School desegregation case challenging use of NTE for teacher selection under court-ordered plan. | Use of NTE as one factor in employment decisions not improper. Remanded on other grounds. |
| 1972 | Spurlock | Challenge to airline's requirement of college degree and minimum flight hours as discriminatory. | Although disproportionately affecting blacks, the requirements were upheld because they were shown to be job-related, to have "business necessity," and to be justifiable in view of high degree of human risk. |

Table 3: Continued

| Date | Case | Description | Court Decision |
|---|---|---|---|
| 1975 | Tyler | Class action by blacks against Georgia bar exam charging purposeful discrimination and denial of due process and equal protection. | Exam found neither intentionally nor inherently discriminatory and was rationally related to a legitimate state purpose. Lack of review did not deny due process, since right to retake exam satisfied. |
| 1978 | N.E.A. | Constitutional and Title 7 challenge to use of NTE validated for certification of South Carolina teachers on grounds it disproportionately disqualified blacks. | Supreme Court affirmed NTE despite its disproportionate effect on blacks, minus proof of discriminatory intent in its use. |

13

Table 4: Complaints Dealing with Equal Protection that Were Upheld by the Courts

| Date | Case | Description | Court Decision |
|---|---|---|---|
| 1969 | U.S. v. Metal W. | Employment discrimination suit against subjective, pass-fail union journeyman's exam. | Union practices found discriminatory. |
| 1971 | Griggs | Class action against diploma requirement and employment testing as discriminatory. | Supreme Court found power company's requirements preserved past discrimination and were not shown to be job-related. |
| 1972 | W.A.C. Organ. | Challenge to unvalidated fire department exam which adversely affected blacks. | Fire department failed to demonstrate job-relatedness of exam after a prima facie showing of discriminatory effect. Court enjoined use of exam. |
| 1972 | Castro | Unvalidated police entrance exam challenged as discriminatory. | Massachusetts Civil Service Commission was unable to demonstrate exam's job-relatedness after a showing of discriminatory impact. Court ordered a new exam and relief in eligibility classification. |

Table 4: Continued

| Date | Case | Description | Court Decision |
|---|---|---|---|
| 1972 | Armstead | School district's use of GRE score, AA teaching certificate or master's degree for employment decisions contested as discriminatory. | Court ruled GRE requirement was not reasonably related to the purpose for which it was intended, thereby denying equal protection. School Board was given opportunity to justify use of AA certificate and master's degree. |
| 1972 | Anderson | Class action by nonminority employees charging reverse discrimination in hiring and promotion to administrative positions as a result of quotas for minority employment. | Quota was held unconstitutional, denying equal protection to nonminorities. |
| 1973 | U.S. v. Georgia Power | Employment discrimination action challenging testing program. | Court rejected post-testing validity study claiming exam job-related because it did not meet EEOC Guidelines. |

Table 4: Continued

| Date | Case | Description | Court Decision |
|---|---|---|---|
| 1973 | Vulcan | Class action against New York City Fire Department's employment practices and civil service exam violating equal protection. | Exam had discriminatory effect and was not sufficiently job-related. Use of eligibility list enjoined. |
| 1974 | Wade | Employment discrimination class action suit focusing upon a job evaluation procedure brought against an agricultural and home economics extension service. | The unvalidated, subjective evaluation procedure had an adverse racial impact. Court forbade continued use. |
| 1974 | N.A.A.C.P. | Challenge to Massachusetts Civil Service exam for hiring firemen, unvalidated prior to use. | Court rejected findings of concurrent validity study as insignificant, finding exam not job-related and cut-off score arbitrary. Exam enjoined and color-conscious relief ordered. |
| 1975 | Albemarle | Class action by black employees against their employer and labor union to enjoin discriminatory employment practices and testing program. | Supreme Court found Albemarle's validation study defective and held testing program not job-related. Injunction against use of test vacated since company policy altered during appeal. |

Table 4: Continued

| Date | Case | Description | Court Decision |
|---|---|---|---|
| 1975 | Kirkland | Class action against New York State Correctional Services promotion exam as discriminatory. | Exam was not job-related and failed to meet legally required standards. Court enjoined use and ordered development of a new exam. |
| 1977 | Chance | Class action charging exams for selecting and licensing New York City school supervisors resulted in de facto discrimination against blacks and Puerto Ricans. | After a prima facie showing of discrimination, Board of Examiners failed to justify necessity of exam. Resolution of which of two replacement supervisory selection plans, both of which were federally acceptable, was to be instituted was left up to state courts. |

## Summary of Findings

The body of cases surveyed during this investigation were those charging violation of Fourteenth Amendment guarantees of due process and equal protection, as well as Title VII class action suits. Primary concerns arising from a review of these cases are the necessity of proving the instrument used for evaluation and employment decisions job-related, as well as insuring that the instrument is validated in accordance with federal regulations.

The Supreme Court rulings in *Griggs v. Duke Power* (28 L.Ed. 2d 158) and *Albemarle v. Moody* (45 L.Ed. 2d 280) are landmarks in the matter of employment discrimination focusing upon test validation. These Title VII cases established significant job-relatedness as the primary criterion which employers must meet in setting their evaluation standards. Further, these cases established that where a discriminatory effect results from use of an instrument, it must be validated.

Test construction should follow a thorough job analysis. As it was put in *Kirkland v. New York State Department of Correctional Services* (320 F.2d 420), "insufficient spadework usually results in a poor garden . . ." The court in *Vulcan Society v. Civil Service Commission of the City of New York* (490 F.2d 387) concurred, maintaining that the poor quality of test preparation was considered to reflect badly upon the quality of the exam in that case, and increased the burden upon defendants of showing job-relatedness.

Statutory jurisdiction in the area of employment discrimination centers around Title VII of the Civil Rights Act of 1964 (42 U.S.C. Sec.

2000). This act forbids the discriminatory restriction of employment opportunities by private employers, labor unions, and employment agencies on the basis of race, color, religion, sex, or national origin. The scope of this legislation was extended to include state and local employees, educational institutions and employers in interstate commerce with fifteen or more employees by the Equal Employment Opportunity Act of 1972.

Enforcement of Title VII is divided between the Equal Employment Opportunity Commission (EEOC) for the private sector and state and local government, the U.S. Civil Service Commission (USCSC) for the federal government, and the Librarian of Congress for the Library of Congress.

Constitutional jurisdiction for discrimination suits revolves around the Thirteenth and Fourteenth Amendments and their applicable statutory counterparts, 42 U.S.C. Secs. 1981-1983. The Thirteenth Amendment, enacted in 1865, prohibited "slavery and involuntary servitude," and the courts have ruled that discrimination on the basis of race, color, and national origin falls within the purview of this amendment. The Fourteenth Amendment, enacted in 1868, required that all state and local governments must extend "equal protection of the laws" to all people. This amendment prohibits discrimination individually or by class in both employment and public service.

The provisions of these amendments were incorporated into several civil rights laws. The Civil Rights Acts of 1866, 1870, and 1871 (42 U.S.C. Secs. 1981-1983) guaranteed the universality of equal rights and equal protection. For a state to infringe upon these fundamental, protected individual rights, it must be able to show a "compelling

governmental interest" which requires abridgment, whereas in cases not involving such basic individual rights the government's burden would be less rigorous, requiring only that a "rational government interest" be proven. These statutes have been used by public employees in discrimination suits prior to the extension of Title VII to state and local governments in 1972, as well as by those complainants whose time allowances under Title VII have elapsed.

One result of Title VII of the Civil Rights Act of 1964 upon the conduct of employment discrimination suits was the facilitation of class action suits,* in which only a statistical showing of disproportionate effect of a selection device upon any protected group (race, color, religion, sex, or national origin) is sufficient to prove a prima facie case of discrimination and thereby shift the burden of proving the job-relatedness of the device to the employer. Upon such a showing, the employer must demonstrate that no other less discriminatory criteria would suffice to achieve his goal. The employer's intent under Title VII is immaterial--it is the effect of his actions that is of judicial concern.

Although the complainant in a Title VII case can show differential effect by other means than just a prima facie statistical showing, the latter appears to be the easier and more frequent strategy. Other means for demonstrating differential effect include showing the discriminatory content, nature, or application of the instrument. However, often these

---

*According to William F. Gardner in "The Development of the Substantive Principles of Title VII Law: The Defendent's View," 26 ALABAMA LAW REVIEW 5, most Title VII cases at the district court level are class actions.

latter strategies are secondary and succeed a statistical showing of differential effect. As Judith Bartnoff explains:

> To the extent that a differential impact cannot be established by statistics alone, other evidence derived from an examination of the content and application of specific criteria and selection mechanisms is indicated. But if statistics alone establish a differential impact, then specific criteria need be examined, if at all, only in connection with the employer's attempt to justify his practices.*

Whereas proof of discriminatory *intent* on the part of the employer is required in suits brought under constitutional provisions, the focus of Title VII suits is upon the discriminatory *effects* of employment practices. As the Supreme Court indicated in *Griggs*, ". . . Good intent or absence of discriminatory intent does not redeem employment procedures or testing mechanisms that operate as 'built-in headwinds' for minority groups and are unrelated to measuring job capability." Business necessity is the yardstick for judging questionable practices.

In *Washington v. Davis* (48 L.Ed. 2d 597), the case in which the distinction between Title VII and constitutional cases was drawn, the Supreme Court held that in a constitutional case ". . . it does not follow that a law or official act is unconstitutional *solely* because it has a racially disproportionate impact regardless of whether it reflects a racially discriminatory purpose." Further, in *NAACP v. Beecher* (504 F.2d 1017) the U.S. Court of Appeals (1st Circuit) noted that an employee selection device may disqualify a disproportionate number of minority candidates and still not be discriminatory in the constitutional sense. While proof of

---

*Judith Bartnoff, "Title VII and Employment Discrimination in 'Upper Level' Jobs," in *Winning an Equal Employment Opportunity Case: Prosecution and Defense of Discrimination Actions*, edited by New York Law Journal, 1974, p. 12. Reprinted from 73 COLUMBIA LAW REVIEW 1614.

disproportionate effect is not central to a constitutional case, it is essential to the plaintiff's ability to initiate a Title VII case.

The Supreme Court recently addressed the issue of whether proof of discrimination requires showing discriminatory intent, or whether evidence of discriminatory effect alone is sufficient in a challenge to a teacher certification system in *NEA v. South Carolina* (54 L.Ed. 775). This case, brought on Fourteenth Amendment as well as Title VII grounds, involved South Carolina's use of a minimum score for certification, a requirement instituted after a validation study was completed by the Educational Testing Service (ETS), designer of NTE. The state realized the traditional disparity in minority passage rates of the NTE at the time the requirement was adopted. Despite this disparate impact, however, South Carolina's validated NTE requirement survived challenge, with the Court ruling no discriminatory intent was shown and NTE was rationally related to the purpose for which it was used.

The distinguishing factor in *NEA v. South Carolina*, however, was the fact that NTE had been validated specifically for use in South Carolina's school system, and a cutoff score was derived consistent with this objective. Attempts by school systems to use both NTE and GRE as a means of evaluating in-service teachers without this prior "spade-work" have failed; e.g., *GAE v. Nix* (407 F. Supp. 1102) and *Armstead v. Starkville* (461 F.2d 276). Particularly in cases where institution of an unvalidated NTE requirement paralleled court-ordered desegregation, as in *Baker v. Columbus Municipal Separate School District* (329 F. Supp. 706) and *Walston v. County School Board of Nansemond County, Va.* (492 F.2d 919), such requirement is legally vulnerable to charges of intentional discrimination.

Although the burden of justifying actions may shift quickly to the defendant in a Title VII case upon a prima facie showing of discrimination, the plaintiff cannot successfully bypass this essential first step. In *U.S. v. H.K. Porter* (296 F. Supp. 50), a case in which the government unsuccessfully attempted to disallow an employment test on the basis that it was not "professionally validated," as the government interpreted it to be required under EEOC Guidelines,* the court held that in order to find a test discriminatory, there must be some *evidence* of a discriminatory result. A similar attempt by a teacher in a constitutional case claiming an unvalidated teacher rating form "created an impermissible classification of all teachers deemed incompetent as a result of its use" was unsuccessful in *Beachamp v. Davis* (550 F.2d 959). The court in this case ruled that such a claim was "dubious at best as a matter of law, since there is no allegation that use of the evaluation form disproportionately affected members of any identifiable group . . ."

In assessing the adequacy of the validation strategy employed in *Nansemond*, Judge Hoffman in the District Court proceeding noted that ". . . no study has shown any correlation between any score and an ultimately effective in-service teacher; that is, teacher ability to increase learning capacity of students." Thus, the NTE lacked predictive validity (the method preferred by EEOC), but as Judge Hoffman added, ". . . *No* test currently measures this relationship; nor is it likely that any test could be developed to accomplish this goal." The lack of a universal definition of an "effective teacher," the subjective qualities required of teachers, and the inability to isolate a teacher's effect upon a class

---

*EEOC and FEA Guidelines are not in accord on this point. Although FEA demands test validation only in instances where adverse effect results, EEOC appears to require it of all tests. The courts, however, have not always upheld EEOC's position, as indicated *infra*.

from other variables were cited as reasons for the impossibility of achieving predictive validity. Requiring predictive validation was also considered unreasonable, since such a requirement could adversely affect the children's education by exposing them to low-scoring teachers. Finding the requirement of predictive validation unreasonable and technically infeasible, the NTE was held to be acceptable on the basis of content validity.

Ultimately, however, Judge Hoffman's decision in *Nansemond* was reversed by the U.S. Court of Appeals (4th Circuit) (492 F.2d 919). The Court of Appeals held the "reasonable measure" standard of *Western Addition Community Organization v. Alioto* (340 F. Supp. 1351) was incorrectly applied in this case, maintaining the more rigorous "demonstrable relationship" test of *Griggs* should have been employed. "The NTE, as used by the Board, does not purport to measure the teacher's actual knowledge of the subject matter assigned to be taught or his performance in the classroom, but places primary emphasis on general education and professional education," the court stated. Thus, the NTE fell into the category of the "broad and general testing devices" prohibited by *Griggs*.

The Court of Appeals in *Nansemond* did not rule on the issue of the NTE's content validity, finding that the prerequisites for allowing use of a content valid instrument under EEOC Guidelines were not met. While content validity is permissible "where a well developed test, consisting of suitable samples of essential knowledge, skills or behaviors composing the job in question are posed in the examination," the court stated: ". . . We do not find the necessary prerequisites for such a finding by the District Court since no such samples, job analyses or validation

studies were provided or conducted. Nor did any expert testify that the questions posed tested the existence of the knowledge essential to performance in the teaching positions."

As to the District Court's contention that "no study has shown any correlation between any score and an ultimately effective in-service teacher," the Court of Appeals stated, "If this be true, the NTE cut-off score of 500 is patently arbitrary and discriminatory." Further, the fact that the county superintendent realized that blacks were less likely to score as high as whites on NTE was magnified still more by the institution of NTE as the *sine qua non*.

Yet, although Nansemond's use of NTE was found to be discriminatory and the court ordered reinstatement with backpay for teachers dismissed solely because of NTE and reevaluation of teachers dismissed for cause, the court did *not* rule that NTE could not be used under any circumstances:

> What we hold today is not to say that the NTE cannot be utilized under any circumstances as one possible type of "objective criteria" under the rule in *Chambers* . . . However, it cannot be used as a tool of discrimination; it cannot be used without proper validation studies and job analyses under the rule in *Moody* . . . Nor can it be administered capriciously in derogation of the guidelines promulgated by the ETS so as to have a racially disparate impact. It may well be that when properly and fairly applied in appropriate situations, the NTE could qualify as having the "demonstrable relationship," as required by *Griggs* . . . so essential to ensure equal employment opportunity.

Commenting on the requirement of predictive validation, Michael Rebell states:

> Predictive validation, that is, empirical demonstration that credentialling or licensing requirements are directly correlated to specific on-the-job performance needs, is a standard which public employers do not yet seem capable of achieving, certainly not in the teacher education area. The courts have indicated that, for the moment, until predictive validation

is "feasible," they will generally accept content validation in licensing cases.*

The courts have recognized the need for a thorough job analysis, however, and Rebell maintains that this is the key to acceptable content validation.

Projecting future developments in the area of validation, William Gardner holds that while the courts have been reasonable and realistic in their approach to validation studies conducted in accordance with the EEOC Guidelines as well as those that represent provisional compliance, ". . . the defense cannot rest on precedent alone, for it is certain that the plaintiffs will attack the validation studies offered by defendants on the basis of every conceivable defect which they might allegedly contain."** Potential attacks might center around claims that the criteria of performance represented in the validation study do not constitute "major or critical work behaviors," as required by the guidelines, or that either the cultural anxieties of those being tested or the attitudes of those giving the test adversely affected scores. Another challenge might arise from a claim that performance ratings employed in a study did not accurately reflect the actual performance of the subjects.

Despite some fairly optimistic developments in due process/equal protection cases--particularly that of *NEA v. South Carolina*--an employer's burden in an employment discrimination action is still formidable. While

---

*Michael A. Rebell, "The Law, the Courts, and Teacher Credentialling Reform" in *Licensing and Accreditation in Education*, edited by Betty Levitov (Lincoln, Nebraska: University of Nebraska Printing and Duplicating Service, 1976), p. 14.

**William F. Gardner, "The Development of the Substantive Principles of Title VII Law: The Defendant's View," 26 ALABAMA LAW REVIEW 74-75.

theoretically it is the complainant's task to prove that discrimination has occurred, Gardner maintains that realistically ". . . the more accurate view . . . is that the defendant has the burden of proving that it is not guilty of discrimination." Because of the burden frequently imposed on defendants of rebutting a prima facie case, Gardner maintains that ". . . any idea that a defendant can prevail if the plaintiff does not prove the allegations of discrimination contained in the complaint is definitely ill-advised in the trial of Title VII actions."*

## Observations

Upon reviewing the findings of this investigation, several observations can be made which, after further investigation, may be utilized in determining needed changes in the content or procedures for administering the TPAI. These initial observations include:

1. The state has the right to establish and change requirements for teacher certification, and to delegate this authority to a state department of education.
2. To be legally secure an employment instrument should be in accord with the Equal Employment Opportunity Commission's "Guidelines on Employee Selection Procedures."
3. To be legally secure an employment instrument should be in accord with the Federal Executive Agency's "Guidelines on Employee Selection Procedures."
4. Evaluation instruments should meet the American Psychological

---

*Gardner, pp. 8-9.

Association's *Standards for Education and Psychological Tests* (1974) if they are to pass a court challenge.

5. Where an instrument can be shown to have a disproportionate effect upon any protected class, validation is legally required.
6. In meeting the burden of proving the job-relatedness of a test or evaluation instrument, the findings of a job analysis figure prominently.
7. In a class action suit brought under Title VII, if no discriminatory effect can be demonstrated upon any of the protected groups under that statute, validation is not legally required.
8. In employment testing not only the instrument itself, but also the procedure by which it is administered is vulnerable to challenge.
9. The scoring system used in employment testing should be reasonably related to job performance. Cutoff scores cannot be arbitrary.
10. The provision of an adequate method of review as well as the opportunity to be retested after an applicant has gained additional training greatly strengthens the equitableness of an examination procedure.

## Plans for Continuation

Although this study to date has succeeded in providing a general understanding of the field of employment testing and has introduced some of the potential legal issues raised by a system of performance-based teacher certification, a number of questions still remain unanswered.

To a large extent, the failure to produce clear definitions of acceptable validation techniques results from a lack of consensus among experts in a very complex field. In addition, the lack of legal precedent on very technical aspects of validation procedures reflects the fact that most litigation to date has not been conducted on this level. Many of the tests challenged had never been validated. Also the fact that employers have frequently agreed to alter their practices under the terms of conciliation agreements or consent decrees to avoid costly litigation has probably subtracted from developments in this area. Hopefully such cases broaching more technical aspects of testing will begin to materialize, facilitating a more comprehensive understanding of "acceptable" testing practice.

In addition to seeking more detailed precedents on testing, an attempt will also be made to follow up appeals in cases already documented, since it is not unusual for this type of litigation to span a decade in the appeals process and to be reversed several times along the way.

## SELECTED BIBLIOGRAPHY

Court Cases


Albermarle Paper Co. v. Moody, 422 U.S. 405, 95 S.Ct. 2362, 45 L.Ed. 280 (1975).

Allen v. City of Mobile, 466 F.2d 122 (5th Cir. 1972), certiorari denied 412 U.S. 909 (1973).

Anderson v. San Francisco Unified School District, 357 F. Supp. 248 (N.D. Calif. 1972).

Armstead v. Starkville Municipal Separate School District, 461 F.2d 276 (5th Cir. 1972).

Baker v. Columbus Municipal Separate School, 329 F. Supp. 706 (N.D. Miss. 1971), affirmed 462 F.2d 1112 (5th Cir. 1972).

Beachamp v. Davis, 550 F.2d 959 (4th Cir. 1977).

Boston Chapter, N.A.A.C.P., Inc. v. Beecher, 504 F.2d 1017 (1st Cir. 1975), certiorari denied 43 L.Ed. 775 (1975).

Brito v. Zia Company, 478 F.2d 1200 (10th Cir. 1973).

Castro v. Beecher, 459 F.2d 725 (1st Cir. 1972).

Chance v. Board of Examiners, 561 F.2d 1079 (2nd Cir. 1977).

Georgia Association of Educators v. Nix, 407 F. Supp. 1102 (N.D. Ga. 1976).

Griggs v. Duke Power Company, 401 U.S. 424, 91 S.Ct. 849, 28 L.Ed. 2d 158 (1971).

Guthrie v. Taylor, 185 S.E. 2d 193 (N.C. 1971).

Hodge v. Stegall, 242 P.2d 720 (Okla. 1952), rehearing denied.

Kirkland v. New York State Department of Correctional Services, 520 F.2d 420 (2nd Cir. 1975), rehearing and rehearing en banc denied.

Lee v. Macon County Board of Education, 463 F.2d 1174 (5th Cir. 1972).

Longarzo v. Anker, 358 N.Y. Supp. 2d 936 (N.Y. 1974).

Marrs v. Matthews, 270 S.W. 586 (Tex. Ct. Civ. Ap. 1925).

Matteson v. State Board of Education, 136 P.2d 120 (Calif. Ct. App. 1943).

Meffert v. State Board of Medical Registration and Examination, 72 P. 247 (Kans. 1903).

National Education Association v. South Carolina, 434 U.S. 1026, 54 L.Ed. 775 (1978).

Spurlock v. United Airlines, Inc., 475 F. 2d 216 (10th Cir. 1972).

Stone v. Fritts, 82 N.E. 792 (Ind. 1907).

Tyler v. Vickery, 417 F.2d 1089 (5th Cir. 1975), rehearing en banc denied and certiorari denied 521 F.2d 815 (1976).

U.S. v. Georgia Power Co., 474 F.2d 906 (5th Cir. 1973), rehearing denied.

U.S. v. H.K. Porter Co., 296 F. Supp. 40 (N.D. Ala. 1968), appeal docketed, No. 27703, 5th Cir., Mar. 14, 1969.

U.S. v. Sheet Metal Workers Local No. 36, 416 F.2d 123 (8th Cir. 1969).

Vulcan Society v. Civil Service Commission of the City of New York, 490 F.2d 387 (2nd Cir. 1973).

Wade v. Mississippi Cooperative Extension Service, 372 F. Supp. 126 (N.D. Miss. 1974).

Walston v. County School Board of Nansemond County, Virginia, 492 F.2d 919 (4th Cir. 1974).

Washington v. Davis, 426 U.S. 229, 96 S. Ct. 2040, 48 L.Ed. 2d 597 (1976).

Western Addition Community Organization v. Alioto, 340 F. Supp. 1351 (N.D. Calif. 1972).

Guidelines and Standards

American Psychological Association, *Standards for Educational and Psychological Tests* (Wash., D.C.: APA, 1974).

Equal Employment Opportunity Commission, "Guidelines on Employee Selection Procedures," *Federal Register*, 1 August 1970, pp. 12333-12336.

Federal Executive Agency, "Guidelines on Employee Selection Procedures," *Federal Register*, 23 November 1976, pp. 51734, 51752-51759.

## APPENDIX A

### Equal Employment Opportunity Commission's "Guidelines on Employee Selection Procedures"

The EEOC Guidelines (originally issued in 1966 and last revised in 1970) were issued in response to the Commission's concern with the increased use of testing for employment decisions, often as the sole criterion for hiring or transferring, and the possible discriminatory results of such employment testing. The Guidelines were developed in "the belief that properly validated and standardized employee selection procedures can significantly contribute to the implementation of nondiscriminatory personnel policies, as required by Title VII."

According to the Guidelines, the term "test" is defined as any paper-and-pencil or performance measure used as a basis for any employment decision. Discrimination is considered to be "the use of any test which adversely affects hiring, promotion, transfer or any other employment or membership opportunity of classes protected by Title VII . . . unless: (a) the test has been validated and evidences a high degree of utility . . . and (b) the person giving or acting upon the results of the particular test can demonstrate that alternate suitable hiring, transfer or promotion procedures are unavailable for use." Protected groups under this legislation are those identifiable by race, color, religion, sex and national origin.

Validation consists of empirical evidence demonstrating that the test is "predictive of or significantly correlated with important elements of work behavior which comprise or are relevant to the job or jobs

for which candidates are being evaluated." The American Psychological Association's *Standards for Education and Psychological Tests and Manuals* is cited as a guide for acceptable validation procedures. While the EEOC appears to favor criterion-related validity as defined by APA, either content or construct validity is permissible where criterion-related validity is not feasible, although either alternate method should be supported by job analysis information demonstrating its relevance. "Well-developed tests that consist of suitable samples of the essential knowledge, skills or behaviors composing the job in question" may be supported by evidence of content validity alone.

A number of minimum standards for validation are specified. In conducting concurrent validation studies upon incumbent employees, the sample must be representative of minority applicants. Tests must be administered and scored under controlled and standardized conditions, and test score security must be safeguarded. The work behaviors which the test is intended to predict should be fully described, and should be crucial as shown by a careful job analysis. Supervisory ratings should be cautiously developed and examined for bias. Moreover, all criteria should be examined for any factor which could unfairly depress minority scores. In the course of differential validation, ". . . where a test is valid for two groups but one group characteristically obtains higher test scores than the other without a corresponding difference in job performance, cutoff scores must be set so as to predict the same probability of job success in both groups."

Where technically feasible a test should be separately validated for *each* minority group for which it is used. "Technical feasibility" is

defined as "having or obtaining a sufficient number of minority individuals to achieve findings of statistical and practical significance, the opportunity to obtain unbiased job performance criteria, etc." The employer claiming absence of technical feasibility must be able to demonstrate the same. Any disproportionate effect of the test upon minorities must be related to job performance.

In determining the "utility" of a test, both statistical significance and practical significance should be considered. The relationship between the test and at least one relevant criterion must be statistically significant--i.e., the probability of the relationship having occurred by chance should be no more than one in twenty. The test should also evidence practical significance; i.e., the larger the number of applicants placed on jobs through use of the test, the greater the need for demonstrating a high degree of relationship between the test and the criterion. Moreover, the larger the proportion of applicants selected by other means who become successful employees, the higher the relationship should be between the test and criterion. In jobs where the economic and human risks of hiring an unqualified person are minimal versus rejecting a qualified applicant, the higher the relationship should be. Selection standards also must be applied uniformly--there can be no disparate treatment of individuals or protected classes under Title VII. Further, opportunity for retesting should be provided for earlier failures who have acquired more experience and/or training.

Evidence that tests are not being used in a discriminatory manner should be kept in readily available form for inspection. Validation results should be presented in graphical and statistical form, and

average scores for all relevant subgroups shown. Copies of tests, manuals and administration instructions, scoring information and minimum cutoff score which are not available through normal commercial channels should be included with validation evidence. In addition, "whenever statistical adjustments are made in validity results for less than perfect reliability or for restriction of score range in the test or the criterion, or both, the supporting evidence from the validation study must be presented in detail."

## APPENDIX B

### Federal Executive Agency's "Guidelines on Employee Selection Procedures"

The FEA Guidelines, issued in 1976, were developed in an attempt to provide a uniform framework for employee selection standards within the federal government and to provide more up-to-date technical information in this area.

Validation is required by these guidelines only where a selection device has an adverse effect upon one or more of the protected groups under Title VII. "Adverse impact" is defined as a selection rate for any group which is less than 4/5 or 80% of the selection rate for the highest scoring group. A selection rate above 4/5 will generally be considered to constitute evidence that there is no adverse effect. However, "smaller differences in selection rate may nevertheless be considered to constitute adverse impact where they are significant in both statistical and practical terms." Where adverse effect occurs, the individual components of the test should be examined closely for evidence of disproportionate effect.

Three different methods of validation are recognized--criterion-related, content and construct validity as defined by the APA. Validation should proceed from a job analysis. The measures of the job or criteria which the instrument is designed to reflect should constitute critical work behaviors. Statistical significance between the selection procedure and job performance should represent no more than a one in

twenty possibility of having occurred by chance. Care should be taken to avoid bias in the selection and application of measures of performance, particularly in supervisory ratings and subjective evaluations. The validation sample should be representative of the applicant population and adequate in size. Allowance is made for use of an unvalidated procedure provided substantial evidence can be cited in accord with the Guidelines, and studies are in progress to produce the required evidence of validity.

Where adverse impact occurs, an investigation of fairness should be made. The Guidelines state that "an investigation of fairness of a selection procedure depends on both evidence of validity and the manner in which the selection procedure is to be used in a particular employment context."

Some procedural concerns include the requirement that the instrument must be administered and scored under standardized conditions. There can be no disparate treatment of members of racial, sex or ethnic groups. Cutoff scores must be established to reflect reasonable work proficiency.

Data should be collected and kept readily available for inspection showing the effect of a selection procedure upon specific sex, racial and ethnic categories. Separate classifications should include: Blacks (Negroes), American Indians (including Alaskan Natives), Asians (including Pacific Islanders), Hispanics (including persons of Mexican, Puerto Rican, Cuban, Central or South American, or other Spanish origin or culture regardless of race), Whites (Caucasians) other than Hispanics, and totals. Caution should be exercised in assuring that this information is used for research purposes concerning adverse impact, however, and not for employment decisions.

## APPENDIX C

### APA *Standards for Educational and Psychological Tests*

The 1974 APA *Standards* are a revision of the original 1966 edition and an attempt to present up-to-date guidelines on assessment procedures or devices both for test users and test developers.

According to the *Standards*, each test should have an accompanying manual so that a test user may ascertain the test's usefulness for his purposes. The characteristics the test is intended to measure, as well as its rationale, development, technical features, administration and application should be clearly described in such a manual. Evidence of the test's validity and reliability, including data on the standard error of measurement and reliability coefficients, and the manner in which these measurements were derived should be presented along with clearly defined procedures for scoring. The absence of any necessary evidence should be noted where applicable. In addition, the manual should assist users in making correct test interpretations and should warn against common misuses. Qualifications for individuals who will administer the test and interpret test scores should also be specified.

Test norms should be published in the manual and should refer to a clearly defined population, employing percentiles or standard scores in the reporting of such reference groups. Derivation of scales from normative data should be clearly described. Measures of central tendency and variability should also be presented in the manual. Where appropriate, the use of local norms may be suggested. Cutoff scores should be directly related to the overall purpose of the test.

At appropriate intervals, the test manual should be revised. Such revision is necessary "whenever changing conditions of use or new research data make any statements in the manual incorrect or misleading." Where a test is intended to be used over a long period of time, a mechanism should be established for gathering data for continued research.

Three validation techniques are recognized by APA: criterion-related (including both predictive and concurrent approaches), content and construct validity. As much information on validation efforts as possible should be presented in the manual, so that the user can critically evaluate the test. Evidence of validity should be presented for each type of inference for which test use is proposed. Validity statements should not be based on the original sample, but should be cross-validated. While necessary evidence of validity may be found in a well-documented manual, the test user is responsible for "marshalling the evidence in support of his claims of validity and reliability."

Where criterion-related validation is attempted, criteria measures should be completely and accurately described, and the adequacy of each criterion commented upon. Criterion measures should be studied for evidence of validity. Limitations upon the generalizability of validity information should be provided.

In content validation where a representative performance sample is cited in a universe of situations, the manual should clearly define the universe and the sampling procedures. By contrast, in construct validation where an interpretation of scores is proposed to measure a theoretical variable, the proposed interpretation should be fully defined and distinguished from other possible interpretations.

Procedures followed in conducting the validation effort (sampling, collection of data, etc.) should conform to the standards of recommended testing procedure, and should be described in sufficient detail for the reader to assess the test's usefulness.

Wherever possible, validity studies should be performed separately for subsamples identifiable by national origin, race, sex, or religious affiliation, and results of such studies should be reported in the manual.

A test user should be familiar not only with literature relevant to the tests he uses, but with the overall area of testing and related problems. He should be conversant with the principles of measurement, validation and testing research, and should realize the limitations of test interpretation.

The selection of or development of a test should reflect clearly formulated goals and hypotheses. Choice of an existing test should be based on the compatability of the test's research and development with its intended use. In selecting or developing a test, one should attempt to minimize "tester effect" and maximize tester reliability.

Great care should be taken to insure that bias is avoided in test selection, administration and interpretation. The test user should examine test items for bias. Tests should be administered under standardized conditions, and test security should be maintained.

Institutional users should review tests they administer periodically to take into account current developments.